# Modular Method of Detection, Localization and Counting of Mutliple-Taxon Pollen Apertures Using Bag of Words

Gildardo Lozano-Vega, Yannick Benezeth, Franck Marzani, Frank Boochs

# Modular Method of Detection, Localization and Counting of Mutliple-Taxon Pollen Apertures Using Bag of Words

Gildardo Lozano-Vega[1,2], Yannick Benezeth[2], Franck Marzani[2], and Frank Boochs[1]

[1] i3mainz, Fachhochschule Mainz, Lucy-Hillebrand-Strasse 2, 55128 Mainz, Germany
gildardo.lozano@fh-mainz.de
[2] Le2i, Université de Bourgogne, B.P. 47870, 21078 Dijon Cedex, France

**Abstract.** Accurate recognition of airborne pollen taxa is crucial for understanding and treating allergic diseases, which affect an important proportion of the world population. Modern computer vision techniques enables the detection of discriminant characteristics. Apertures is one of these characteristic that has been little explored up to now. In this paper, a flexible method of detection, localization and counting of apertures of different pollen taxa with varying appearances is proposed. Apertures are described based by primitive images following the Bag-of-Words strategy. A confidence map is estimated based on the classification of sampled regions. The method is designed to be extended modularly to new aperture types employing the same algorithm by building individual classifiers. The method was evaluation on the top 5 allergenic pollen taxa in Germany and robustness to unseen particles was verified.

**Keywords:** pattern recognition, LBP, bag of words, apertures, palynology.

## 1 Introduction

Allergic diseases due to airborne pollen affect considerably the world population. The correct estimation of the pollen concentration in the surrounding of individuals becomes important for the study of allergies. Particularly, the recognition of the taxa at genus level is relevant for the diagnosis, medication and prevention of allergic diseases. Traditional methods consist of counting manually the pollen particles under the microscope from airborne samples. These methods are time consuming and costly, limiting considerably the frequency and the extension of the region under analysis. Furthermore, this type of methods are highly susceptible to human error and inconsistency.

The recognition of pollen taxa is a complex task due to the high diversity of flowering plants. On one hand, pollen particles of the same taxonomic family usually share many similarities even when they belong to different genus branches (inter-class similarity). On the other hand, pollen are naturally formed particles

whose characteristics present variations even in the same genus branch (intra-class variation).

Palynological literature describes morphological and aerobiological pollen characteristics, which are employed by specialists for the recognition of taxa [1][2]. Average size, shape, equatorial outline, aperture type and amount, exine and intine structure, and type of ornamentation are among the main morphological characteristics, although not all can be precisely measured through common microscopes. Aerobiological characteristics refer to the period of pollination for each taxon and can be obtained by historic data from previous years. This information indicates the time slots when high levels of concentration of pollen are present in the environment.

The type and amount of apertures are particularly characteristic of the pollen taxa. Sometimes, they are the key factor to discriminate between two similar genus. Located at the pollen external wall, apertures are morphological distinctive regions which can be observed often with a common microscope due to its size. The appearance of the apertures varies according to the pollen taxon. Moreover, the orientation and focal plane of the particle with respect to the observer can also change their appearance considerably, and even make the aperture difficult to observe. Examples of allergenic pollen taxa are shown in Figure 1. In some of them, apertures are not visible at all.


**Fig. 1.** Examples of most important allergic pollen in Germany. Apertures of *Alder*, *Birch* and *Hazel* pollen are highlighted by a green circle. Note the different appearances for the same pollen type. Although *Mugwort* and *Grass* also own apertures, they are rarely visible on the image due their size or position on the particle.

Common strategies for image-based classification of pollen employ global features like shape, texture and color; or local invariants [3][4][5][6][7]. Despite its importance, few works on detection of apertures exist. Boucher *et al.* developed specific algorithms according to the pollen type and to the point of view in order to detect the aperture, reticulum and cytoplasm of different pollen taxa [5]. Although no much detail is provided, different techniques were employed for each case such as thresholding, Laplacian of Gaussian, and region segmentation. The identified characteristic was then described by means of color, shape and size. Also oriented to the detection of aperture and colpus, Chen *et al.* [6] created different algorithms for *Birch* and *Mugwort* pollen. Hough transform and thresholding were used for detection of circular apertures. For apertures on the particle

image border, they analyzed the intensity profile through the polar transform and applied a template matching strategy and frequency spectrum analysis to detect specific aperture types. Similar to Boucher *et al.*, multiple algorithms were created for the diversity of pollen types, appearances and points of views. These approaches are inflexible if learning more different apertures is needed, requiring the development of new specific algorithms. In both cases, the aperture information was fed to a greater classification system for the recognition of taxa. Therefore, no individual evaluation of the aperture detection was shown, limiting the performance comparison among approaches.

We propose a different approach for the detection, localization and counting of apertures for different pollen types which overcomes the variability due to changes of appearance. The method relies on the Bag-of-Words (BOW) approach for the description of apertures in terms of primitive images. Then, multiple region images are sampled and classified to eventually create by voting, a likelihood map of the presence of apertures on the pollen image. The amount of apertures can be used eventually together with global features for the classification of pollen taxa. The proposed approach is able to learn new varieties of aperture by following a unique methodology, eliminating the need of designing new algorithms or methods. This flexibility is particularly important for growing the system with apertures with multiple appearances.

The rest of the paper is organized as follows: In section 2, the description of the apertures, detection, localization and counting methods are explained in detail. In section 3, experiments on region configuration and detection results are presented. Conclusions about the proposed approach and future work are given in section 4.

## 2 Recognition and Localization of Apertures

In this section, the methodology to represent apertures and to train the classification models is described. Additionally, the process of analysis of an unseen pollen particle for the detection of apertures is explained. First, a general overview of the method is presented in order to acquire the central idea of the proposal, and then the detailed procedure is described.

### 2.1 Outline of the proposed method

The proposed method starts describing the apertures by their content of primitive subimages. These primitive subimages, which form a visual codebook, are represented by a local descriptor and clustered following a similar procedure to BOW. Based on the codebook representation, a classification model learns to discriminate aperture regions form the rest of the particle. Following the same methodology, individual classifiers are created oriented to the different pollen taxa. In the evaluation stage, multiple image regions inside the particle are classified using the proposed method. Regions classified as *aperture* vote positively while the rest negatively. The votes are pixel-wise averaged on a confidence map

which indicates the likelihood of the presence of an aperture. Finally, a local maxima algorithm allows to determine the position and amount of the apertures. The overall scheme of the method can be seen in Fig. 2. The application of individual classifiers in the same manner yields to the detection of the different apertures types, which combined results in the definitive detected aperture list.

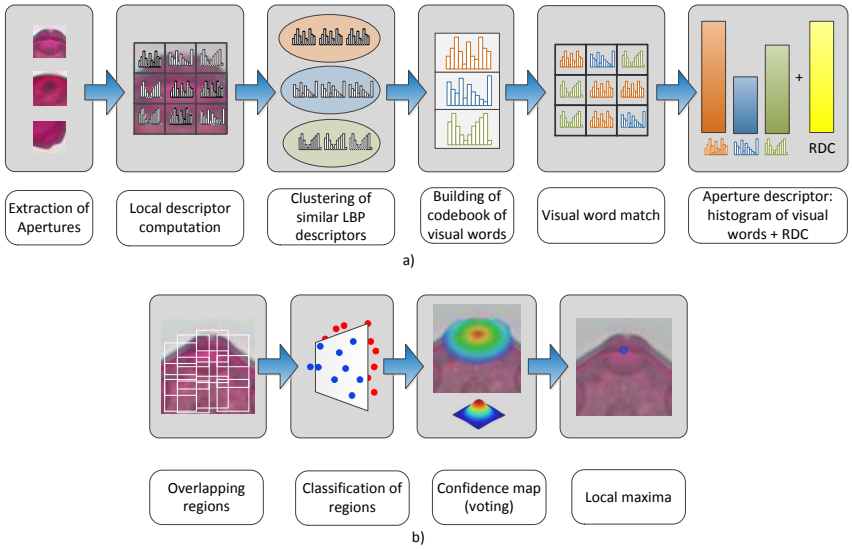



**Fig. 2.** Method of representation of apertures and detection on unknown particles. a) Aperture descriptor: LBP codes computed on a dense grid of aperture regions are clustered to create the codebook of visual words. Then, apertures are described by the histogram of visual words combined with the relative distance to the centroid ($RDC$) as spatial information. b) Detection scheme of an aperture: Multiple overlapping regions are evaluated and classified to create the confidence map from which the aperture position is estimated.

### 2.2 Description of apertures based on Bag of Words

The foundation of the proposed strategy is the recognition of visual patterns present in the pollen apertures, that allow them to be described in terms of primitive subimages. This idea originated from the Bag-of-Words strategy (also known as Bag of Features) for text categorization, where a codebook is created from words contained in the analyzed text [8]. Then, classification is possible by observing the frequency of each word on the text. The visual counterpart

and its extensions has been typically applied to object recognition [9][10]. However, to our knowledge, the proposed method is the first application of BOW to microscopic objects.

First, different areas of the aperture are represented by a corresponding local descriptor. Although typical BOW indicates that such areas are determined by salient keypoints, an alternative is to compute a dense grid of fixed patches instead [11]. This technique has shown to be efficient by extracting more information from the whole analyzed object with an implied increase of computation. Because apertures regions are not expected bigger than 40 pixels, dense computation of patches is feasible. The current method employes a square grid of square patches, whose size can be determined according to the expected aperture size. Since the grid is set over the aperture region, most of the patches are informative. An example of a grid of 4x4 patches of 8 pixels by side (denoted as 4-8 region) is shown in Fig. 3.

**Fig. 3.** Example of a 4x4 grid employed for the computation of local descriptors. The patch size is a square of 8 pixels.

**LBP as Local Descriptor.** A local descriptor is needed to extract the visual pattern from each patch of the grid. Typically, descriptors are chosen based on its performance together with a keypoint detector, for example SIFT. However, since the proposed method does not need a keypoint detector, only local descriptors are concerned. A comprehensive survey of local descriptors can be found in [12]. In previous experiments, local jets [13], a differential descriptor suitable for shapes, shown to be acceptable for the classification of aperture regions [14]. In the current work, Local Binary Patterns (LBP) operator is proposed and tested as local descriptor of the patches due to the gray-level and rotation invariance, in addition to the simple computation and short representation.

LBP is a simple but efficient operator that condenses the spatial pixel arrangement in a small image vicinity [15]. LBP has been widely employed as texture descriptor in complex tasks, for example, in face recognition [16]. Considering a central pixel $\{x,y\}$ as the reference position, its grey level is compared one-to-one to the levels of neighbor locations. The individual comparison results are counter clockwise concatenated as bits into a binary code, which take the value 0 if the reference pixel has a greater gray level than the corresponding

neighbor pixel, and 1 otherwise. This strategy makes the operator invariant to monotonic gray-level change on the image. The idea can be extended to a neighborhood of radius $r$ and limiting the number of compared surrounding pixels (and consequently the bit length of the code) to $p$ pixels yielding to a code denoted as $LBP_{p,r}$. A graphical explanation of the obtaining of a LBP code is shown in Fig. 4.

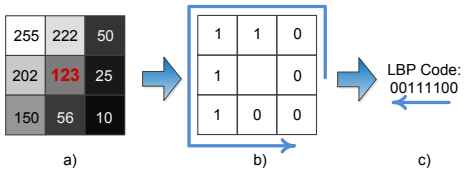



**Fig. 4.** Example of computation of the LBP code from a single pixel. a) Values of the gray levels of the 8 neighbors are shown overlapped to each pixel location. The value of the reference pixel is highlighted in red at the center. b) Individual binary results of the comparison to the reference value. c) Final concatenated LBP code.

Due the bit order, similar pixel arrangements result in different codes when the area is rotated. Ojala *et al.* noticed this fact and reduced the possible LBP codes by grouping rotated patterns in a unique code, providing additionally rotation invariance [17]. Furthermore, they defined a set of uniform patterns ($u_2$) in which the transitions among bits (0/1 change) on the LBP code is not oftener than 2. The rest of the patterns were grouped in a single code. Ojala *et al.* showed that the rotation-invariant uniform patterns, denoted as $LBP^{riu2}_{p,r}$, are still effective since they represent on average 90% of the textural information of an image. The final amount of possible codes is reduced considerably, for example, a $LBP^{riu2}_{8,r}$ has only 10 possible binary codes.

In the proposed method, the LBP descriptor of an image patch is obtained by estimating the normalized LBP code histogram of all the pixels, except those on the border according to the radius $r$. The $LBP^{riu2}_{8,1}$ 10-value descriptor was chosen for the proposed method due to the reduced size of the patches. Pollen images are expected to be scanned at the same scale and therefore, strong scale invariance is not required. Moreover, the change of the size of the apertures can be neglected for pollen of the same type.

**Creation of the Codebook of Visual Words.** Due to the natural variability and the observation point, apertures of the same type cannot show identical LBP patch descriptors. In order to classify the same visual patterns having similar LBP representation , it is necessary to create groups of similar descriptors. Each created group is the equivalent of visual words of the BOW. Assigning a word of the codebook to each LBP patch descriptor is a process that requires

clustering techniques. In this work k-means clustering is employed due to its general effectivity on multiple clustering problems and to its flexibility to chose the number of clusters $k$, i.e., the size of the codebook. On early work, it was demonstrated that adding spatial information of the patches improves the classification performance of the regions [14]. For this reason the coded location of the patch was incorporated to the final patch descriptor. Taking advantage of the source grid where patches are regularly spaced, the position of the patch inside the grid is simply estimated as a number ranging from 0 to 1. Once created, the cluster model will serve to match visual words to unseen image patches. In the experimental section, variation of the size $k$ of the codebook is evaluated with respect to the performance of classification of regions. In Fig. 5, examples of visual words are depicted with some instances of their image patches.

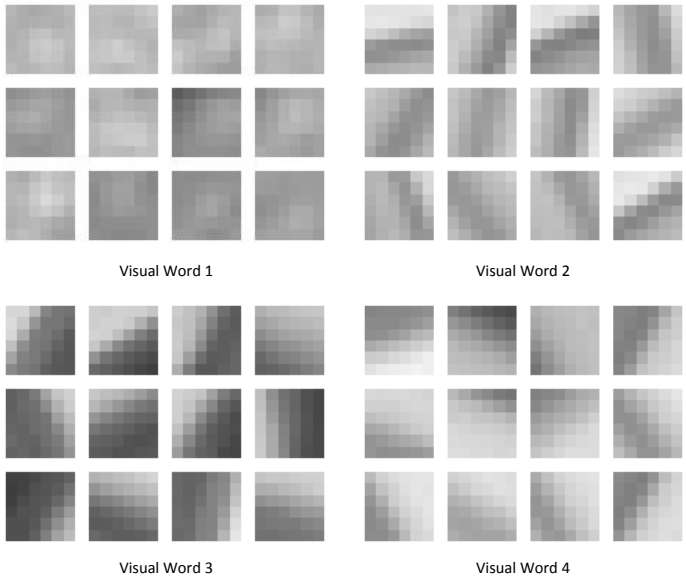



**Fig. 5.** Example of 6-pixel patches grouped according to the assigned visual word. Note the similarity of the patches of the same visual words. Patches are amplified for visualization purposes.

**Aperture descriptor** The aperture descriptor is composed by 2 elements: the content of visual words and the spatial information relative to the particle. The first element is given by the computation of the normalized histogram of visual words of the aperture. This means that each bin of the histogram describes the

frequency of each of primitive subimage inside the aperture region. Similar apertures will have a very similar histogram distribution. The codebook is directly related to the histogram: the bigger the codebook, the larger the histogram. The second element of the description consists of the relative distance of the aperture with respect to the centroid of the particle. Apertures change their appearance according to the orientation of the particle with respect to the focal plane, which also affects the position of the aperture on the image. By considering this position, information about the orientation of the apertures is also provided to the classifier. The position is computed as the relative distance from the central point of the aperture region to the centroid of the particle ($RDC$), where a distance 0 corresponds to the centroid itself and a distance 1 corresponds to the outermost point of the particle. Fig. 6 shows an example of the $RDC$ computation of some regions inside a particle.

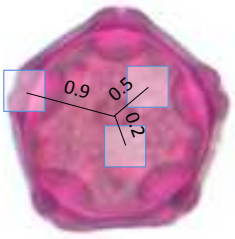



**Fig. 6.** Example of some regions at different position inside the particle and its relative distance to the centroid $RDC$.

### 2.3 Individual Classifiers

An important characteristic of the proposed method is the possibility of easily add a new type of aperture or a set of aperture appearances from a single pollen type, using the same algorithm and without the need of classifying everything already learnt over again. The task is split modularly into individual classifiers for each added aperture type and they eventually will be combined into an aggregate detection. This means that when a new aperture needs to be added to the detector, only the new module has to be created and combined with the existing modules of the rest of the apertures. Since the total of detected apertures is the sum of the detections of the individual classifiers, the problem has optimal substructure property and, therefore, the optimal solution can be obtained by finding independently the optimal solution for each of the individual classifiers.

It is possible to find different appearances of the same aperture type, and therefore, a classifier able to handle such complexity is needed. In a previous work, it was shown that the Support Vector Machine (SVM) method was suitable

for classification of apertures with multiple appearances [14]. SVM is chosen due to proven results in binary classification based on mapping data into a higher dimension to enable linear separation [18].

Therefore, an individual classifier is trained as a binary SVM classifier, where aperture regions are regarded positive and regions not belonging to an aperture are regarded as negative. In order to guarantee that the classifier is able to discriminate both classes among all possible pollen types, samples of negative regions from all pollen types are included in the training dataset. The mapping kernel employs the radial basis function (RBF), and the parameter $\gamma$ and the penalty $C$ are tuned in a 3-fold cross-validation.

### 2.4 Detection of apertures from an unknown image

The SVM classifier delivers a confidence value for each analyzed region. The value can be interpreted as the level of belonging of the region to the positive class. A final binary classification is obtained by comparing the confidence value to a given decision threshold ($d\_thr$) in order to obtain the decision for a positive (aperture) or negative class (not an aperture). Multiple regions covering an specific area can vote, according to their assigned class, to reinforce the belief of the existence or absence of an aperture in the particle. In the proposed method, several overlapping regions covering the whole particle are classified and vote to build an aperture confidence map of the particle. High confidence areas of this map represent detected apertures.

The method does not require mandatorily prior segmentation or knowledge of the particle to be evaluated. However, we have added a fast trace of the contour in order to avoid the analysis of useless regions outside the particle contour. Similarly to the thresholding of Chen *et al.*, our proposal employs Otsu thresholding for binarizing the image which is nonparametric, unsupervised and fast. The principle of the method is to find a threshold that minimizes the intra-class gray-level variance of the background and the foreground pixels. From the resulting binary image, the contour is estimated with Suzuki's method [19] using the OpenCV implementation [20]. Excellent performance can be obtained, provided that the source image exhibits a good contrast and there is no debris stuck to the pollen.

A set of overlapping square regions of equal size are sampled from the particle's inner area defined by the contour. Then, the aperture descriptor from section 2.2 is computed on each sampled region of the particle. Classification confidence is obtained after applying the individual classifier and the decision threshold ( $d\_thr$ ) is employed in order to have a binary decision.

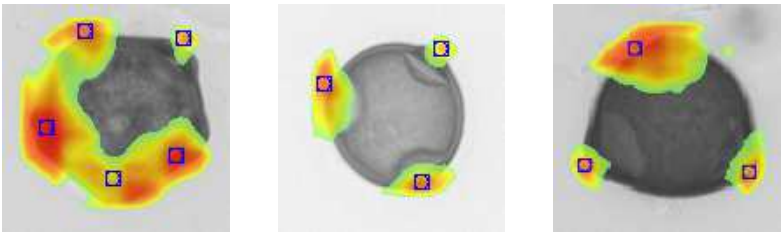
In order to merge the aperture information of the sampled regions, a confidence map is created the same size that the particle image. Each classified region contributes with its vote pixel-wisely to build the map. Regions with positive class votes with value 1 and negative with value 0. The final value of the confidence map is the averaged contribution of all the regions votes expressed as follows in Eq. 1:

$$CM(x,y) = \frac{\sum_{i=1}^{n} V^i}{n} \tag{1}$$

where $CM(x,y)$ is the value of confidence map for the pixel located in $(x,y)$, $V^i$ is the vote of the region $i$, and $n$ is the total of classified regions that contain the pixel in $(x,y)$. The greater $n$, the less impact of a single region on the final value, making the estimation of the map more resilient to misclassification of the regions. The resulting dense confidence map can be interpreted as the likelihood of the presence of apertures on the particle. Therefore, high values on the map are regarded as detected apertures.

The procedure to detect apertures is to find local maxima on the confidence map. Because low confidence areas can cause false detections due to low local maxima, only locations whose confidence value is above a confidence threshold (*c_thr*) are analyzed. Before searching local maxima, the confidence map is smoothed with a gaussian filter of one quarter the size of the regions, in order avoid noise from irregularities due to region sampling. Finally, the $(x,y)$ positions of the local maxima are regarded as potential detection of apertures.

A final procedure is applied before considering the potential detected apertures as definite. If more than one detection are closer each other than a given detection distance (*d_dist*), the detections are regarded of the same aperture, and therefore merged. In such case, the definitive single detected aperture is located at the middle point of the former potential apertures. This technique helps to reduce multiple detection on the same aperture. Detected apertures from individual classifiers are consolidated together to yield to the definite list of detected apertures of the system. Fig. 7 shows examples of the confidence map of different particles and the position of the detected apertures.

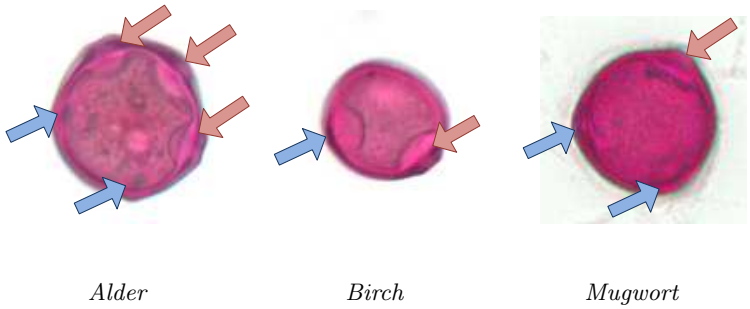


*Alder* *Birch* *Mugwort*

**Fig. 7.** Example of confidence maps overlapped over the original pollen image *Alder* (left), *Birch* (center) and *Hazel* (right) pollen. Warm colors (red-orange) indicates high values of the map, regarding a high probability of the presence of an aperture. On the contrary, cold colors (blue-green) indicate low values of the map. The position of the detected aperture after finding local maxima is shown by a blue mark.

### 2.5 Evaluation of the detection

The apertures change gradually their appearance depending on the observation point, from barely visible to clear and well defined. The proposed method is expected to be trained with enough defined apertures in order to reduce the false positive rate. However, it is still possible to detect unexpected apertures that are not well defined due to their similarity to the well defined apertures. Consequently the ground truth of apertures, which is used to evaluate the performance of the detection, is split into two sets. The ground truth set (GT) groups those enough defined apertures for which the method was trained and whose misdetections are penalized in the evaluation. The unexpected set (UE) groups those not well defined apertures which were not expected to be detected by the algorithm and whose misdetections are not penalized in the evaluation. Examples of both sets are shown in Fig.8



*Alder* *Birch* *Mugwort*

**Fig. 8.** Examples of the aperture sets of *Alder*, *Birch* and *Mugwort* used for evaluation. Well defined apertures marked by a red arrow pointing left belong to the ground truth set (GT). Not well defined apertures marked by a blue arrow pointing right belong to the unexpected set (UE).

The evaluation of the detection consists of quantizing the match between detected apertures and the ground truth of the location of apertures. A distance matrix is employed to compare the detections (DT) on one side, to the group defined by GT and UE on the other side. The euclidean distance is computed one-to-one between the apertures of both compared groups. An example of this matrix is shown in Fig. 9a. Each DT aperture is matched to the closest GT or UE aperture as long as the distance is below an evaluation threshold (*ev_dist*) [Fig. 9b]. A true positive (*hit*) is regarded for each GT aperture that was matched to at least one detection. A false negative (*miss*) is regarded for each GT aperture that was not matched to any detection. An unexpected true positive (*unexpected hit*) is regarded for each UE aperture that was matched to at least one detection. A false positive (*false alarm*) is regarded for each detection that not matched to any GT or UE aperture and for each multiple detection (2 or more detections of

the same aperture) [Fig. 9c]. True positive of UE apertures are not considered for the total of true positives, and their false negative are neither penalized since the method did not learn to detected them; and the statistics are computed just to gain additional insight into the performance.

| | GT1 | GT2 | GT3 | UE1 |
|---|---|---|---|---|
| DT1 | 30 | **5** | 33 | 50 |
| DT2 | 40 | 32 | **8** | 31 |
| DT3 | 33 | 41 | **15** | 42 |
| DT4 | 36 | 37 | 44 | **12** |

a) distance matrix

| | GT1 | GT2 | GT3 | UE1 |
|---|---|---|---|---|
| DT1 | | 1 | | |
| DT2 | | | 1 | |
| DT3 | | | 1 | |
| DT4 | | | | 1 |

b) matrix of matches

| | Apertures | Total |
|---|---|---|
| True Positives | GT2 GT3 | 2 |
| False Negative | GT1 | 1 |
| False Positive | DT3 | 1 |
| Unexpected True Positives | DT4 | 1 |

c) statistics

**Fig. 9.** Example of evaluation of 4 detected apertures (DT) against 3 ground truth (GT) and 1 unexpected (UE) apertures. The distance matrix show the euclidean distance between the groups (a). Minimum distances are highlighted in bold. If the distance is below an $ev\_dist = 27$ pixels, the DT aperture is matched as 1 in the matrix of matches (b). Finally, evaluation statistics are extracted form the matches (c).

With these statistics; true positive, false negative and false positive; is possible to compute a curve similar to the Receiver Operating Characteristic (ROC) that compares the true positive rate (*recall*) versus the false positive rate (*fall-out*) by varying *c_thr*. This kind of plot is useful to visualize the performance of the system under different conditions. If the specific cost function is known for the problem, a point (or segment) on the ROC curve becomes optimal [21]. Contrary to a typical ROC curve, where the threshold variation applied to a classification confidence value directly affects the trade-off between true positives and false positives, the computation of the ROC curve for of the aperture detection is result of a more complex process after the application of the *c_thr* (local maxima, refinement and matching) and the trace of the curve could be not always smooth. Additionally, since there is no a ground truth for the negative class (actually it is represented for all the regions of the particle that are not an aperture), the total of detected apertures is considered for the computation of the false positive rate, instead of the typical total of negatives. This change makes more sensible the ROC curve to a small number of detected apertures, affecting the left-bottom side of the curve.

## 3 Experiments and Evaluation

The proposed method was evaluated with focus on the 5 most allergenic pollen types in Germany: *Alder, Birch, Hazel, Mugwort* and *Grass*. Although the 5 types possess apertures, the tested system was trained to detect apertures of three of them: *Alder, Birch* and *Hazel*. *Mugwort* apertures are hardly visible since they are usually hidden in a furrow. *Grass* apertures are small compared to the size of the pollen, fact that makes them rarely visible in most of the focal planes. For this reason, the amount of *Grass* apertures in our database was vey limited for training. However, the incorporation of *Mugwort* and *Grass* samples in the experiments is important in order to evaluate the robusteness to multiple pollen and to simulate similar conditions to a real-world application.

Images of pollen were scanned with a brightfield microscope Keyence BZ-9000 with magnification 40x and a resolution of 0.26 microns/pixel. Particles were stained with magenta dye. The following amount of particles for each taxa were employed: *Alder* 115, *Birch* 136, *Hazel* 52, *Mugwort* 120 and *Grass* 132. The total of ground truth labeled apertures is 541.

The presented methodology of aperture description, detection, localization, counting and evaluation was applied on the database and 3 individual classifier were built for *Alder, Birch* and *Hazel*. In order to test robustness to unseen data, a 4-fold cross validation was employed on the input dataset, meaning a training/testing ratio of 0.75/0.25.

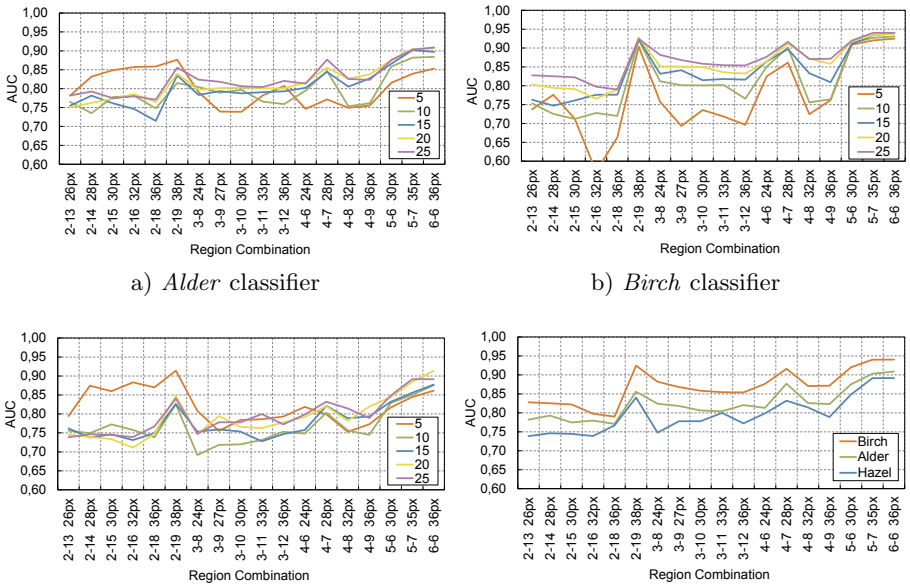
### 3.1 Evaluation of region parameters and codebook size

For the first evaluation step, the correct configuration for patch size, grid size and number of visual words of the codebook needs to be selected based on the performance of the SVM classification of the regions.

Since the variation of the decision threshold *d_thr* is evaluated together with the confidence threhshold *c_thr* in the matching phase, the selected combination has to evidence the best performance at different levels of *d_thr* instead of just a fixed value. The Area Under the Curve (AUC) of the ROC curve of the region classification was employed as comparison metric in order to chose the best combination, robust to different levels of the *d_thr*. For this metric, a perfect performance for all *d_thr* obtains an AUC = 1, and an average random perfomance obtains an AUC = 0.5.

Different region sizes, ranging from 24 to 38 pixels, along with their combination of grid size-patch size and number of visual words were tested. The evaluation was repeated for each individual classifier. Fig. 10(a-c) shows the behavior for each classifier of AUC for different combinations. For the 3 individual classifiers, 25 visual words is generally the best and most consistent selection of codebook size. Although more visual words may improve the performance, it also increases the computation time, especially for a great amount of regions. Fig. 10d compares AUC for the 3 classifiers using 25 visual words. Best results seem related to bigger region sizes. Furthermore, for greater grid sizes (5 to 6), performance seems more stable. Particularly, the 6-6 region (36 pixels) shows

better performance and this was the configuration used for the aperture detector. Similar combinations, like the 5-7 region could have been good election too.




a) *Alder* classifier b) *Birch* classifier

c) *Hazel* classifier d) 25 visual words

**Fig. 10.** AUC performance of region SVM classification for indivudual classifiers: a) *Alder)*, b)*Birch* and c)*Hazel*. d) A comparative of the same classifiers just for 25 visual words.

### 3.2 Evaluation of classification performance

Using a 6-6 region configuration, the pollen dataset was sampled, ranging from 30 to 60 overlapping regions per particle, depending on the size. A SVM model was created from the training set for each fold of the cross validation scheme. The Radial Basis Function (RBF) was employed as kernel, the SVM parameter $\gamma$ and penalty $C$ were optimized using a grid approach with AUC of the ROC as performance metric. A classification confidence value for each sample was computed after applying the SVM model to the testing set. The learning process was reapeated for each individual classifer.

In order to select the values of *d_thr* and *c_thr*, each classifier was evaluated on the detection of apertures and a family of ROC curves was plotted. Fig. 11 shows example of the curves. Distances *d_dist* and *ev_dist* were set experimentally to 27 and 35 respectively. Assuming a balanced cost function, where the cost of missing a detection is equal to the cost of a false positive, the optimal performance is

found in the closest point to the upper left corner of the ROC, where the optimal *d_thr* and *c_thr* can be determined. Due to the imbalanced datasets, the *d_thr* is offset towards zero, being the optimal value relatively low.

As a final step, the output of the optimized individual classifiers were combined to create the final list of detected apertures from the 5-taxon dataset. The final list was evaluated against the ground truth of all the pollen types and the following performance statistics were computed as average of the 4-fold evaluation with confidence level of 0.95. For apertures of the 3 learnt pollen types, joint performance showed a true positive rate of $0.78 \pm 0.03$ , a false positive rate of $0.23 \pm 0.05$, and a precision of $0.76 \pm 0.05$. Considering all the 5 pollen types, the true positive rate remains $0.78 \pm 0.03$, but the false positive rate moved to $0.29 \pm 0.06$, and the precision to $0.70 \pm 0.06$. The difference is caused to false positive detections on *Mugwort* and *Grass*. There is no evidence that other detectors of aperture have been evaluated, and therefore there is no baseline for comparison. However, the resuts are promising considering the high difficulty of learning 3 different aperture types with multiple appearances and adding 2 pollen types which apertures are not easily visible.

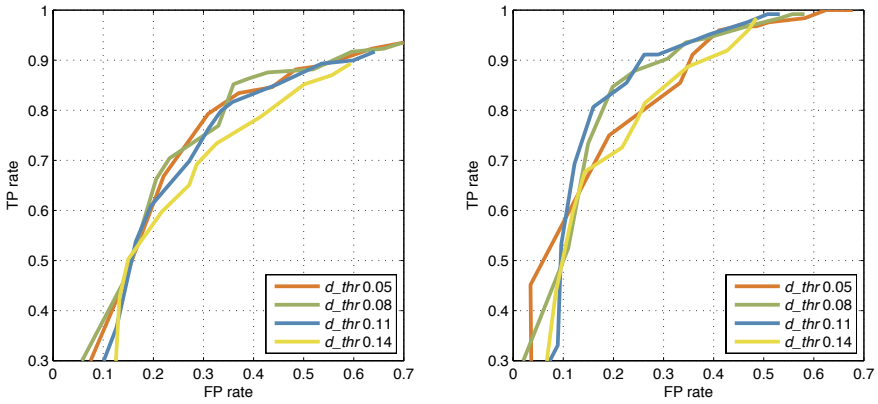
## 4 Conclusions and future work

We have proposed a flexible image-based method to detect, localize and count apertures of pollen. The method is capable of learning independently apertures from different pollen types employing individual classifiers and with the same algorithm. Furthermore, the BOW approach makes the method robust to changes in appearance of the same aperture type. Additionally, the proposed region sampling strategy allows to adjust the configuration for each aperture type.

Experiments confirmed suitability of the method in similar conditions to a real-world application by adding extra pollen taxa with no visible apertures. The cross-validation evaluation showed that the method is robust to unseen particles. The method succeeded under total lack of previous knowledge about location and amount of the apertures and under challenging conditions such as intra-class variation, inter-class similarity, multiple input taxa and analysis on a single focal plane.

Because there are not performance results of the existing methods, there is no baseline for comparison among methods. However, aperture counts, computed using the proposed method, have been recently employed in a pollen classification system together with shape, texture and size features [22]. The performance was comparable to that of Chen *et al.* [6] and superior to Boucher *et al.* [5], who have also employed aperture detection on their pollen classification experiments.

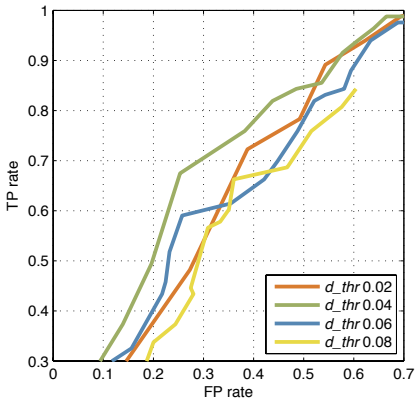
Since apertures change their appearance when observed at different focal planes, next step is the extension of the method for multiple images gathered at these planes. One path is to apply the method to multiple focal planes independently. The advantage would be the confirmation of apertures already detected and the detection of missed apertures. A second path is the application of the lo-




*Alder* classifier




*Birch* classifier




*Hazel* classifier

**Fig. 11.** ROC curves of evaluation on aperture detection for indivudual classifiers: a) *Alder)* b)*Birch* c)*Hazel*. For clarity, only curves for fold 4 are shown as an example. Different points of the curve are created by variating *c_thr*. Variation of *d_thr* is represented by different curves.

cal descriptor involving multiple focal planes at the same time, maybe employing a volume LBP operator instead, which describes the volumetric aperture.

**Acknowledgements.** The authors are grateful to Celeste Chudyk, Yann Ryann and Morad Larhriq for scanning and preparing the pollen datasets and to the Max Plank Institute for Chemistry for permitting the use of their facilities.

## References

1. Erdtman,G.: An Introduction To Pollen Analysis. Chronica Botanica Company, U.S.A (1943)
2. Hesse, M., Halbritter, H., Weber, M., Buchner, R., Frosch- Radivo, A., Ulrich, S.: Pollen Terminology. An illustrated handbook. Springer, Austria (2009)
3. Rodríguez-Damián, M., Cernadas, E., Formella, A., Gonzlez, A.: Automatic identification and classification of pollen of the urticaceae family. In: Proceedings of Advanced Concepts for Intelligent Vision Systems (ACIVS 2003). 38–45 (2003)
4. Li, P., Treloar, W. J., Flenley, J. R., Empson, L.: Towards automation of palynology 2: the use of texture measures and neural network analysis for automated identification of optical images of pollen grains. Journal of Quaternary Science, 19 (8). 755–762 (2004)
5. Boucher, A., Hidalgo, P. J., Thonnat, M., Belmonte, J., Galan, C., Bonton, P., Tomczak, R.: Development of a semi-automatic system for pollen recognition. Aerobiologia. 18(3), 195–201, Springer Netherlands (2002)
6. Chen, C., Hendriks E.A., Duin R.P., Reiber, J., Hiemstra, P., De Weger, L., Stoel, B.: Feasibility study on automated recognition of allergenic pollen: grass, birch and mugwort. Aerobiologia. 22, 275-284 (2006)
7. Ronneberger, O., Wang, Q., Burkhardt, H.: 3D invariants with high robustness to local deformations for automated pollen recognition. In: Proceedings of the 29th DAGM conference on Pattern recognition. 425–435 (2007)
8. Csurka, G., Dance, C., Bray, C., Fan, L., and Willamowski, J.: Visual categorization with bags of keypoints. In: Pattern Recognition and Machine Learning in Computer Vision Workshop, ECCV Grenoble. 1–22, France (2004)
9. Wu, J., Tan, W-C., and Rehg, J.M.: Efficient and Effective Visual Codebook Generation Using Additive Kernels. Journal of Machine Learning Research. 12, 3097–3118. Georgia Institute of Technology (2011)
10. López-Sastre, R. J. and Tuytelaars, T. and Acevedo-Rodríguez, F. J. and Maldonado-Bascón, S.: Towards a more discriminative and semantic visual vocabulary. Computer Vision and Image Understanding. 115, 415–425. Elsevier Science Inc., New York, USA, (2011)
11. Nowak,E., Jurie, F., Triggs, B.: Sampling Strategies for Bag-of-Features Image Classification. In Computer Vision ECCV Lecture Notes in Computer Science vol. 3954, 490–503 (2006)
12. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence. 27, 1615–1630 (2005)
13. Schmid, C., Mohr, R.: Local grayvalue invariants for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence. 19, (5)530–535 (1997)
14. Lozano-Vega, G., Benezeth, Y., Marzani, F., Boochs, F.: Classification of Pollen Apertures Using Bag of Words. In A. Petrosino (ed.), ICIAP (1), 712-721, (2013)

15. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. Pattern Recognition vol. 29 (1996)
16. Huang, D., Shan, C., Ardabilian, M., Wang, Y., Chen, L.: Local Binary Patterns and Its Application to Facial Image Analysis: A Survey. Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, Vol.41(6), 765–781 (2011)
17. Ojala, T., Pietikäinen, M., Mäenpää, T.: Gray scale and rotation invariant texture classification with Local Binary Patterns. In: Computer Vision, ECCV Proceedings, Lecture Notes in Computer Science 1842, Springer, 404 – 420 (2000)
18. Byun, H., Lee, S.: Applications of support vector machines for pattern recognition: A survey. In: Pattern Recognition with Support Vector Machines, First International Workshop Canada, Proceedings. 213–236 (2002)
19. Suzuki, S. and Abe, K.: Topological Structural Analysis of Digitized Binary Images by Border Following. CVGIP 30(1). 32–46 (1985)
20. Open Source Computer Vision Library. `http://opencv.org`
21. Provost, F., Fawcett, T.: Robust Classification for Imprecise Environments. Machine Learning, 42(3), 203–231 (2001)
22. Lozano-Vega, G., Benezeth, Y., Marzani, F., Boochs, F.: Analysis of Relevant Features for Pollen Classification. In XX (ed.), AIAI (1), XX-XX, (2014)